

Alternate Assessment Alignment Study
Report to the Alaska Department of Education and
Early Development

Conducted by

Meagan Karvonen
Western Carolina University

and

Patricia Almond
University of Oregon

Commissioned by
Dillard Research Associates

Contact:

Gerald Tindal
Dillard Research Associates
32534 Dillard Road
Eugene, Oregon 97405
1-800-838-3163
geraldt@uoregon.edu

Organization of Alaska Alternate Assessment Alignment Report

Executive Summary

- Introduction
- Reading
- Writing
- Mathematics
- Science

Introduction

- Description of Alternate Assessment
- Description of Experts
- Reliability
- Analysis
- Report Overview

Reading Alignment Results

- ExGLEs to PSGLEs
- Alternate Assessment to ExGLEs

Writing Alignment Results

- ExGLEs to PSGLEs
- Alternate Assessment to ExGLEs

Mathematics Alignment Results

- ExGLEs to PSGLEs
- Alternate Assessment to ExGLEs

Science Alignment Results

- ExGLEs to PSGLEs
- Alternate Assessment to ExGLEs

EXECUTIVE SUMMARY

This study reports findings from an investigation of the alignment of Alaska's 2007 Alternate Assessments (AAs), Extended Reading, Writing, and Mathematics (ExRWM), to Alaska's content standards and grade level expectations. Alaska assesses student performance in reading, writing, and mathematics in grades 3 through 10. Alaska's Extended Grade Level Expectations (ExGLEs) in reading, writing, and mathematics served as the content basis for this study. This report addresses alignment between (1) Alaska's performance standard grade level expectations (PSGLE), and extended grade level expectations; and (2) ExGLEs and Alaska's 2007 alternate assessment tasks/items. An alignment study into the areas of accessibility of the assessments and instruction in academic content linked to the assessments will be produced as a supplement to this report and will be submitted in July 2007.

Also included in this report is an investigation of the alignment of a pilot assessment, Extended Science, to Alaska's content standards and grade level expectations in science. The science alternate assessment is scheduled to become operational in 2007-2008 in grades 4, 8, and 10. Draft ExGLEs in science that provided the basis for the alternate assessment alignment study in science.

Reading Alignment Results

Reading Extended GLEs. The reading ExGLEs replicate the wording of the performance (*content*) standards in the reading PSGLEs. The academic link is strongest for the mid/secondary grade bands. Panel members determined that 45% of the ExGLEs for grades 3/4 were adequately aligned to the PSGLEs. For grades 5/6, 67% of the ExGLEs were adequately or well aligned. The percent adequately to well aligned for grades 7/8 and 9/10 was 95% and 94% respectively. When only the well aligned category is considered, there is a dramatic increase across the grade pairs. Ten percent of the ExGLEs are well aligned in grades 3/4, 33% are well aligned in grades 5/6, 63% in grades 7/8, and 78% in grades 9/10.

Extended Reading Tasks and Items. When examining the *extent of linkage* between the reading alternate assessment items and the ExGLEs, the link was strongest for the lower elementary grades. A number of the extended reading alternate assessment (AA) items were referenced to ExGLEs with a *fine* grain link (the expectations): 72% for grade band 3/4, 51% for grade band 5/6, 37% and 23% for grade bands 7/8 and 9/10 respectively. The fine grain matches were stronger at the elementary grades. Coarse grain matches strengthened the link between the reading items and the ExGLEs in the elementary grades with an added 22% for grade 3/4 and 2% for grade 5/6. When content experts re-examined the items for mid/secondary grades, many items were linked to performance (*content*) standards from another (different) grade band.

Based on the four basic Webb indices, overall the linking was strongest at grade 3/4 and weakest at grade 9/10. It is important to note that all of the Webb statistics were interpreted with the assumption that the reading AA would present the full range of ExGLE coverage. Without specific information about the emphasis for the target content, judgments about observed vs. intended content could not be made.

All of the elementary reading performance (*content*) standards are represented on the reading AA assessment (categorical concurrence). At the mid/secondary level three content strands are represented at grade 7/8 and two strands at grade 9/10. The cognitive demand (DOK) for the majority of reading AA items was at or above the cognitive demand of the corresponding ExGLEs, ranging from 66% of the items at grade 5/6 to 100% of the items at grades 3/4 and 9/10. When examining balance of representation or the extent to which reading AA items are evenly distributed across content standards, overall balance ranged between .43 and .58 for three grade bands. The balance index for grade 5/6 showed very limited balance across content standards (.19). The range-of-knowledge index indicated that the span of knowledge needed by students to respond correctly to the reading AA items was restricted when compared to the span of knowledge expected by the corresponding standards in the ExGLEs. All grade levels were below the 50%, a criterion applied to general assessments.

Reading Recommendations. The reading AA is an academic reading assessment. Overall the link is stronger at the elementary level. It was also noted that the relationship between the cognitive demand in the ExGLEs and the reading AA items was comparable, that the items are distributed across the content to some extent, and that the span of knowledge students need to respond to the items is restricted compared to that expected on the ExGLEs. There are no clear criteria for judging the adequacy of alignment for alternate assessments although Webb has made some recommendations. The U.S. Department of Education indicates that alternate assessments should be related to grade-level content for the grade in which the student is enrolled. The guidance acknowledges that alternate assessments can be “restricted in scope or complexity,” expecting the depth, breadth or complexity of content to be different from the general assessment.

In this context, we make the following observations/recommendations:

- Complete the reading blueprints across the grade bands and indicate the intended breadth, depth, and cognitive demand. Use the crosswalk information from the task/item/point documents to inform the next version.
- Review reading AA items. ExGLEs improve linkage of ‘on’ grade band content, especially at the middle and secondary levels, and increase the number of items on the reading AA that are linked at the fine grain level.
- Both balance of representation and range-of-knowledge correspondence results were below the level recommended for adequate alignment in general assessments. It is not clear that the same criteria apply to alternate assessments judged against alternate achievement standards. Refining the test specifications and improving the on-grade linking between items and expectations is a higher priority at this time.

Writing Alignment Results

Extended GLEs in Writing. The writing ExGLEs replicate the wording of the performance (*content*) standards presented in the writing PSGLEs. The academic link between the ExGLEs and the PSGLEs is relatively strong for the mid/secondary grade bands. Panel members determined that 29% of the ExGLEs for grades 3/4 were adequately aligned to the PSGLEs. For grades 5/6, 40% or 2 ExGLEs were aligned, one of these partially. The percent adequately to well aligned for grades 7/8 and 9/10 was 95% and 94% respectively. Two ExGLEs were rated well aligned, one in grade band 7/8 and one in grade band 9/10. No elementary ExGLEs were found to be well aligned to PSGLEs. There are no ExGLEs for two of the five writing strands, “citing sources” and “using resources.” Having ExGLEs for these strands may be helpful for educators in planning instruction.

Extended Writing Tasks and Items. The majority of writing AA items at all four grade bands were linked to the ExGLEs at the fine grain level: band 3/4, had 88% of items linked at the fine grain, band 5/6 had 100% linked, band 7/8 had 87% linked, and band 9/10 had 76% of items linked. For some writing AA items at some grade bands, not even a coarse grain or other grade band standard was identified. One particular type of item, total word count, was inconsistently linked and challenged panelists. This item triggered a discussion among the panelists about how to best classify the link.

Based on the four Webb indices, overall the linking between the writing AA and the ExGLEs was strongest at the elementary grades, especially grade band 3/4. As with reading, Webb statistics were interpreted with the assumption that the writing AA would present the full range of ExGLE coverage and did not look at observed versus intended content coverage. For grade bands 3/4, 5/6, and 9/10, all of the elementary writing performance (*content*) standards were represented on the writing AA assessment (categorical concurrence). For grade band 7/8, one standard was not represented. There were no items that referenced PSGLE 3.4. The elementary grade bands fared better than the mid/secondary bands on Webb’s more stringent six item criterion for categorical concurrence. The cognitive demand (DOK) for the majority of writing AA items was below the demand represented in the ExGLEs overall but there was a better match for grade band 3/4 where 71% of the items were at or above the demand on the ExGLEs and for band 7/8 where 62% of the items were at or above the demand for the ExGLEs than for the other two bands. When examining balance of representation or the extent to which writing AA items are evenly distributed across content standards, balance ranged between .42 and .75 and applying the Webb criterion for balance, grade band 3/4 met the criterion, bands 5/6 and 9/10 weakly met the criterion, and band 7/8 showed very limited balance. The range-of-knowledge correspondence index indicated that all grade bands met the criteria that at least 50% of objectives had at least one aligned assessment item: band 3/4 had 71%, band 5/6 had 100%, band 7/8 had 60%, and band 9/10 had 67%, indicating that overall the span of knowledge needed by students to respond correctly to the writing AA items was comparable to that expected by the ExGLEs.

Writing Recommendations. The writing AA is an academic writing assessment that is aligned to the ExGLEs. Overall the link is stronger for the elementary level ExGLE grade bands. It was also noted that the relationship between the cognitive demand in the ExGLEs and the writing AA

items was lower for the assessment than the ExGLEs but closer for grade bands 3/4 and 7/8. There are no clear criteria for judging the adequacy of alignment of alternate assessments to grade level content standards although Webb has recommended criteria for general assessments. The U.S. Department of Education indicates that alternate assessments should be related to grade-level content for the grade in which the student is enrolled. The guidance acknowledges that alternate assessments can be “restricted in scope or complexity” and expect the depth, breadth or complexity of content to be different from the general assessment. In this context, we make the following recommendations:

- Complete the writing alternate assessment blueprints across the grade bands and indicate the intended breadth, depth, and cognitive demand. Use the crosswalk information from the task/item/point documents to inform the next version of the blueprints.
- Although “citing sources” and “using resources” are not flagged to be assessed on the state’s general assessment, Alaska might consider providing ExGLEs for these strands as an instructional support for educators of students with significant cognitive disabilities.
- Review writing AA items and ExGLEs to identify areas where stronger linking can be accomplished with content from the ExGLE middle and secondary grade bands.

Mathematics Alignment Results

Mathematics Extended GLEs. The mathematics ExGLEs replicate the wording of the performance (*content*) standards in the mathematics PSGLEs. The academic link between ExGLEs and PSGLEs is strongest for the 7/8 grade band. When a match was identified between an ExGLE and PSGLE it was typically judged as adequate. In the grade band 9/10, about half of the ExGLEs were judged as not having a corresponding PSGLE.

Extended mathematics Tasks and Items. When examining the degree of linking between the mathematics alternate assessment items and the ExGLEs, the link was strongest for the grade bands 5/6 and 7/8 at the fine grain level. There were a large number of items across all grade bands that fell into the measurement strand of mathematics, approximately 60% of the items were linked to measurement and most of these were not linked to a particular ExGLE at the fine grain level. A number of mathematics AA items that fell into the measurement strand did link at the fine grain to an ExGLE in the mid/secondary grade bands, 20 out of 59 at the 7/8 band and 13 out of 63 at the 9/10 grade band. There were four strands with few or no mathematics AA items: estimation/computation, functions & relationships, statistics & probability, and problem solving.

Based on the four basic Webb indices, overall the linking was strongest at grade 7/8 and weakest at grade 9/10, however, these indices were calculated without the fine grain links for measurement items. It is important to note that all of the Webb statistics were interpreted with the assumption that the mathematics AA would present the full range of ExGLE coverage. Without specific information about the emphasis for the target content, judgments about observed vs. intended content could not be made. The cognitive demand (DOK) for the majority of mathematics AA items was at or above the cognitive demand of the corresponding ExGLEs, ranging from 53% of the items at grade 9/10 to 84% of the items at grade band 7/8. When examining balance of representation or the extent to which mathematics AA items are evenly distributed across content standards, overall balance ranged between .54 and .76 for the grade

bands. The range-of-knowledge index indicated that the span of knowledge needed by students to respond correctly to the mathematics AA items was restricted when compared to the span of knowledge expected by the corresponding standards in the ExGLEs. All grade levels were below the 50%, a criterion applied to general assessments. The range-of-knowledge correspondence index indicated that the elementary grade bands met the criteria that at least 50% of objectives had at least one aligned assessment item: band 3/4 had 67% and band 5/6 had 69%.

Mathematics Recommendations. The mathematics AA is an academic mathematics assessment. There are no clear criteria for judging the adequacy of alignment for alternate assessment although Webb has made some recommendations for general assessments. The U.S. Department of Education indicates that alternate assessments should be related to grade-level content for the grade in which the student is enrolled. The guidance acknowledges that alternate assessments can be “restricted in scope or complexity,” expecting the depth, breadth or complexity of content to be different from the general assessment.

In this context, we make the following observations/recommendations:

- Complete the mathematics blueprints across the grade bands and indicate the intended breadth, depth, and cognitive demand. Use the crosswalk information from the task/item/point documents to inform the next version. If the emphasis on the measurement strand across the grade bands is sustained, write a rationale for this emphasis.
- Review mathematics AA items and ExGLEs and work to improve linkage of ‘on’ grade band content, at all grade bands, and increase the number of items on the mathematics AA that are linked at the fine grain level.
- Range-of-knowledge correspondence results for grade band 9/10 was below the 50% criterion level recommended for adequate alignment in general assessments. Although, it is not clear that the same criteria apply to alternate assessments judged against alternate achievement standards, this observation warrants review.
- Refining the test blueprints and improving the on-grade linking between items and expectations is highest priority at this time.

Science Alignment Results

Science Extended GLEs. The academic link between ExGLEs and PSGLEs is relatively strong for all grade bands but is strongest for the 7/8 grade band where between 62% (grade 7) and 71% (grade 8) of the items were adequately to well aligned.

Extended Science Tasks and Items. The emphasis in the science AA is in the Physical, Life, and Earth Sciences. There were five items linked to the Science and Technology strand. Sixty-five percent of the standards were represented by at least one item. Roughly half of the items presented a cognitive demand at or above that represented in the ExGLEs. Science AA items were evenly distributed across the content standards and the assessment items were aligned to multiple objectives, showing an acceptable range-of-knowledge correspondence.

Science Recommendations. The alternate assessment in science was a pilot assessment during the spring of 2007. Alignment results are acceptable at this time. No test blueprint or test

specification was provided for the science alternate assessment for this study. There are no recommendations for the alignment of the science alternate assessment at this time.

Summary

Alaska's alternate assessments are academic assessments for students with significant cognitive disabilities. Based on the findings from this study, there is a relatively strong link from the alternate assessments to the ExGLEs. The ExGLEs, designed to be reduced in breadth and depth from the grade level expectations (PSGLEs), vary in their degree of alignment. They are aligned to the Alaska Standards. Alaska's alternate assessment and extended grade level expectations are still relatively new as a large-scale assessment system; such systems evolve over time and changes can be guided by the findings from this alignment study. The system can and should be further refined to more closely align to grade level content and the areas of content emphasis for the assessments should be further specified in the test blueprints to clarify which content will be emphasized and assessed at which grade bands.

**READING, WRITING, MATHEMATICS, AND SCIENCE
ALTERNATE ASSESSMENT ALIGNMENT STUDY
REPORT TO THE ALASKA DEPARTMENT OF EDUCATION AND EARLY DEVELOPMENT**

INTRODUCTION

This alignment study was conducted on the basis of information obtained about the 2007 Alaska Alternate Assessment System. Alaska administered the Extended Reading, Writing, and Mathematics alternate assessments to students with significant cognitive disabilities in grades 3 through 10 to measure student performance in reading, writing, and mathematics in the spring of 2007. A pilot administration of an alternate assessment in science, Extended Science, was administered for grades 4, 8, and 10 in the spring of 2007.

Nonregulatory guidance has specified that alternate assessments “should be clearly related to grade-level content, although [they] may be restricted in scope or complexity or take the form of introductory or prerequisite skills” (U.S. Department of Education, 2005, p.26). As stated in this regulation, there should be a clear *link* to the content standards for the grade in which the student is enrolled. While this gives states flexibility to determine the scope and breadth of content of alternate assessments, it does not exempt states from designing assessments that measure an academic domain with interpretable results that accurately reflect what the student knows and can do within that academic domain. It is expected there will be some differences in the depth, breadth or complexity of content addressed when alternate assessments judged against alternate achievement standards are compared to grade level expectations.

This alignment study was grounded on Tindal's (2005) adaptation of the Webb model of alignment, which focuses on the content of standards and assessments (including categorical concurrence, range of knowledge, depth of knowledge, and balance of representation). Although Webb's model also considers articulation across the grades, equity, pedagogical implications, and systems applicability, these are not included in the current study. Webb's model (1997, 2002) combines qualitative expert judgments and quantified coding and analysis, yielding a set of statistics for each standard and grade on the degree of alignment between state content standards and state assessments. Trained reviewers individually identify the content standard objectives that match each assessment item and the depth of knowledge required by each objective/benchmark of the content standards being analyzed. Next, reviewers determine the objective/benchmark represented by each item or task on the state assessment and rate the depth of knowledge required for successful completion. Reviewers' ratings are entered into a spreadsheet and are analyzed across reviewers, producing statistics on the four categories (Webb, 2002).

**METHOD FOR STUDYING ALIGNMENT OF ALTERNATE ASSESSMENTS TO
GRADE LEVEL ACADEMIC CONTENT STANDARDS**

Ratings for the alignment study were conducted in Anchorage, Alaska on February 5, 6, and 7, 2007 at the Atwood Building, room 240. Rating of alignment between grade level expectations, extended grade level expectations was conducted by having raters identify which grade level expectation matches the extended grade level expectation and rating the quality of the match.

Alignment of Alaska's alternate assessments with the extended grade level expectations was conducted using an adapted version of Webb's alignment method. Judgments about the content and cognitive demand represented in Alaska's Extended Grade Level Expectations (GLE) and in the Extended Reading, Writing, Mathematics, and Science (ExRWMS) alternate assessments by Alaska educators. Alignment indices were calculated based on those ratings.

For alignment of the Extended GLEs to Performance Standard (content standard) Grade Level Expectations (PSGLEs) and raters were asked to identify the link and rate it on a three point scale. Webb's method requires ratings on two dimensions: content match and cognitive demand, also referred to as complexity or depth of knowledge (DOK). The ratings were applied to the relationship between the ExRWMS tasks and items and the ExGLEs. This two-part approach to the alignment shows the chain of evidence needed to link the alternate assessment to the grade level expectations. First the ExGLEs are referenced to the PSGLEs and then (b) the ExGLEs are linked to ExRWMS, the alternate assessments.

Content matches for a single primary "hit" identified the ExGLE that best matched the ExRWMS task/item. Raters also were asked to identify one or more secondary "hits" as they thought the matches existed. Secondary hits represented ExGLEs that seemed to somewhat reflect the content in the ExRWMS task/item, but not as clearly as the primary hit. Depth of Knowledge (DOK) was measured using a 6-point scale developed by Karvonen, Flowers, Wakeman, & Browder (in press). This scale is an adaptation of Bloom's taxonomy, but extended downward in order to be sensitive to the range of cognitive demand required of students eligible to participate in alternate assessments. Each point on the scale has several associated verbs to help raters distinguish between levels of complexity. The scale is presented in the next section on definitions.

Definitions

1. Categorical Concurrence refers to the degree to which standards and assessments address the same content categories. This criterion is met if both documents display the same or consistent categories of content.
2. Depth-of-Knowledge Consistency, or is the degree to which the depth or complexity of knowledge required by the standards and assessments are in agreement. If the assessment is as demanding conceptually as the expectations standards set for the students, this criterion is met. Depth of Knowledge (DOK) will be judged using a six-point scale developed by Karvonen, Wakeman, Flowers, & Browder (in press), see Table 1, below. This scale is an adaptation of Bloom's taxonomy, but extended downward in order to be sensitive to the range of cognitive demand required of students eligible to participate in alternate assessments. Each point on the scale has several associated verbs to help raters distinguish between levels of complexity.

Table 1. Depth of Knowledge Ratings.

Code	Label	Sample behaviors
1	Attention	touch, look, vocalize, respond, attend, recognize
2	Memorize/Recall	list, describe, identify, state, define, label
3	Performance	perform, demonstrate, follow, choose, count, locate
4	Comprehension	explain, conclude, group, restate, review, translate
5	Application	compute, organize, collect, apply, classify, construct, solve, operate, use, generalize
6	Analysis, Synthesis, Evaluation— <i>continued</i>	take apart, combine, differentiate, divide, isolate, dissect, pattern, analyze, experiment, test, compare, contrast, diagram, compose, predict, extend, hypothesize, create, modify, connect, plan, design, defend, verify, conclude, value, assess, rate, judge, appraise, evaluate
X	Too vague	Behavioral expectation is unclear; can't code

3. Range-of-Knowledge Correspondence is the degree to which the span of knowledge a standard expects of students matches that required to correctly answer the assessment items or activity.

4. Balance of Representation is the extent to which assessment items are evenly dispersed across learning objectives within a standard.

RESOURCE DOCUMENTS FOR ALIGNMENT STUDIES:

- Department of Education & Early Development (February 2006). *Alaska Extended Grade Level Expectations for Students with Significant Cognitive Disabilities, Draft – Version 1*. Retrieved on July 17, 2006 at <http://www.eed.state.ak.us/tls/assessment/AlternateOptional/ExGLEs/ExGLEBookMarch06.pdf>
- Department of Education & Early Development (March 2006). *Alaska Science Extended Grade Level Expectations for Students with Significant Cognitive Disabilities, Draft – Version 1*. Retrieved on January 22, 2006 at <http://www.eed.state.ak.us/tls/assessment/AlternateOptional/ExGLEs/ScienceExGLEBookMarch07.doc>
- *Alaska Standards, 4th Edition* available on the Alaska web site at <http://www.eed.state.ak.us/ContentStandards/>
- 2007 Extended Assessment Scoring Protocols and Student Materials in Reading, Writing, Mathematics, and Science (pilot)

DESCRIPTION OF THE ALASKA ALTERNATE ASSESSMENT SYSTEM

Extended Grade Level Expectations. Alaska's Extended Grade Level Expectations (ExGLEs) have been developed for students with significant cognitive disabilities in grades 3- 10. They are designed to accompany Alaska's existing Performance Standards/Grade Level Expectations (PSGLEs) for grades 3-10. The PSGLEs articulate the skills students need to learn and be able to

do by the end of a given grade level. The ExGLEs extend the targeted content for accessibility by providing prerequisite skills and early entry points for students to access the content. The ExGLEs present extensions in grade pairs, i.e. grades 3-4, 5-6, 7-8, and 9-10. The table below displays the ExGLE counts for each grade pair and each subject area.

Counts of Extended GLE Indicators for Grade Pairs and Subject Areas

Grade Pair ► ▼ Subject	3-4	5-6	7-8	9-10
Reading	20	21	19	18
Writing	7	5	5	6
Mathematics	20	29	32	33
Science	21	21	21	21

The ExGLEs are intended to provide a road map for the development of assessment tasks and items as well as the basis for instruction. Some Extended GLEs have been identified as Local, providing the same look and feel employed in the PSGLEs. The designation “(L)” for local assessment is used in the PSGLEs to indicate that the expectation will not be on a state assessment. The March 2006 version of the ExGLEs mirrors the PSGLEs in the placement of the “(L)” notation.

Example of the Local designation from ExGLE Reading (page 11)

[3] 1.1 –5; [4] 2.1-5 Displaying an understanding of print directionality (L)

Alternate Assessments. The Alaska alternate assessments administered during the spring of 2007 were common cross-grade performance task assessments composed of tasks or testlets containing multiple items. Each assessment was comprised of a Scoring Protocol that presented the directions for administration and the scoring worksheets for scoring student responses accompanied by a set of Student Materials providing test stimuli and manipulatives. The alternate assessments were administered in a one-on-one format with one teacher and one student. The teacher/assessor is part of the assessment instrument and presents the tasks and scores the student responses. The operational alternate assessment system for 2007 included Extended Reading, Writing, and Mathematics. A pilot administration of Extended Science was also administered during the spring of 2007. The Alaska alternate assessments are cousin assessments to extended assessments originally developed in Oregon. A brief description of each subject matter assessment follows. Technical manuals available from Dillard Research Associates, the test vendor, provide a detailed description of the assessments, content domains, plus test and item specifications.

Test Blueprints. The 2007 alternate assessment test blueprints were works in progress at the time of this study. We reviewed draft alternate assessment blueprints in reading, writing, and mathematics (see excerpt below) showing assessable content from the ExGLEs along side Alaska’s Standards Based Assessment Blueprints. We also reviewed crosswalk tables relating alternate assessment tasks and items to ExGLEs (see excerpt below). Crosswalks were available for reading, writing, and mathematics. Neither the draft blueprints nor the crosswalks developed

by DRA had been reviewed or approved by Alaska DEED at the time of the study. Thus, judgments about the alignment of the alternate assessments against standards were made based on the full span of extended GLEs rather than intentionally prioritized content.

Excerpt from Alternate Assessment Reading Blueprint Draft January 2007

Grade	Strand	Performance Standard	SBA Assessable GRADE LEVEL EXPECTATION (GLE)	Reading Distribution of Emphasis on 2005 SBA Operational Test	AA Assessable Extended Grade Level Expectation	AA Assessable Reading Distribution of Emphasis	
3	Fluency	1.3 Read text aloud	NA	0%	NA		
	Word Identification Skills	1.1 Use structural analysis: read words	1.1.1 1.1.3 1.1.4	30%	1.1-1 1.1-2 1.1-3		
	Forming a General Understanding	1.2 Comprehend literal meaning	1.2.1	50%	1.2-1		
		1.4 Retell or restate information	1.4.2		NA		
		1.5 Identify main idea	1.5.1		NA		
		1.6 Follow simple directions	1.6.1		NA		
	Analysis of General Content or Structure	1.7 Identify forms of text	1.7.2	20%	NA		
		1.8 Identify story elements	1.8.1		NA		
		1.9 Express own opinions about text	NA		NA		
		1.10 Make connections	1.10.1 1.10.2		NA		
		1.11 Identify cultural influences	NA		NA		
	Totals for Grade 3				100% (rounded)		
	NA = Not applicable in state assessment, locally assessed Text Types: Literature 60%, Informational 40%						

Crosswalk Table of Extended GLEs with Corresponding Tasks, Items, and Points March 2007

Reading: Table. Extended GLE Measurement by Task, Item and Point Value							
Table 1: Specific							
Grades 3 and 4							
Strand/Attribute	GLE	Task #	Total Tasks	# of Items	# of Points	% of Total Items	% of Total Points
Word Identification Skills	[3] 1.1-1; [4] 2.1-1 Identify signs and symbols	1, 2	2	16	32	19.75	18.18
	[3] 1.1-2; [4] 2.1-2 Identifying at least 10 letter-sound relationships	6	1	8	8	9.8	4.54
	[3] 1.1-3; [4] 2.1-3 Blending at least 5 sounds to make word	9	1	8	16	8.64	7.95
	[3] 1.1-4; [4] 2.1-4 Identify own name in print (ID)	WRITING TASK?					
Forming a General Understanding	[3/4] 1.2, 2.2-1 Identifying detail from a story read aloud using pictures, symbols, or words.	4	1	6	12	7.41	6.81

DESCRIPTION OF CONTENT AND SPECIAL EDUCATION EXPERTS

In reading, writing, and mathematics, each alignment team was made up of three panelists; two content area experts and one special education expert with knowledge and experience with students with significant cognitive disabilities. There were two teams per subject, one for elementary grades and one for middle school and high school. In writing, one content expert was absent and the middle school/high school team was made up of one content expert and one special educator. In science, there was one panel made up of two content area experts and one special education expert with knowledge and experience with students with significant cognitive disabilities. One science content expert had experience with the elementary grades and the other had experience with middle school and high school.

The alignment team had a total of 20 panelists, six for reading, six for mathematics, three for science, and five for writing. Content experts had a range of experience in their content area of 3 to 33 years ($M = 16$ years) and the special educators had a range of experience in special education of 3 to 21 years ($M = 12$ years). Five panelists had less than 10 years of experience in K-12 teaching (3 to 9 years). Of the 13 content experts, nine held masters degrees and one was working toward a masters degree. Of the seven special education experts, four held masters' degrees and one held a doctorate in special education. Five of the special educators and ten of the thirteen content experts provided professional development in their area of expertise. Eight panelists taught in higher education (seven content experts and one special educator).

Panelists were configured in their 3-member teams during the training process. At the beginning of each alignment activity, each team worked together to come to a consensus on the alignment of educational components. When experts disagreed, decision rules were made to ensure consistency. Then the reviewers independently rated a subset of tasks/items and extended grade level expectations, and agreement between raters was examined. When the team members met 90% agreement, each rater was given specific duties. Reliability was checked periodically throughout the alignment activities to ensure consistent judgments.

RELIABILITY OF CODING

Once a set of consensus items had been rated and decision rules created within each team, all remaining data points were coded independently by each team member. Once ratings were completed for one grade pair, a study facilitator compared ratings to identify discrepancies. Tasks, items, and indicators with discrepant ratings were resubmitted to the team for discussion and consensus by the group to obtain a final rating. Ratings were returned when exact agreement was not found between at least two raters. When review of a set of coding forms revealed consistent discrepancies, the consensus process included retraining and refinement of coding rules to improve future consistency. Exact agreement for primary hits and depth of knowledge and the alignment of grade level expectations were calculated on initial ratings.

TRAINING

During training an overview of Alaska's alternate assessments and extended grade level expectations was provided. Panelists learned about the student population and how grade level standards are reduced in breadth and depth. The alignment process was described and panelists were given examples of various standards and tasks that are closely aligned and not closely aligned in reading, writing, mathematics, and science. Teams were instructed about their roles in working together and rating independently after calibration, but serving as resources for each other when unfamiliar terminology was encountered. Each group was provided with an orienting practice exercise.

As part of training, panelists were given a practice exercise to complete individually with a set of standards and tasks. Each panelist completed a form independently coding each task to an extended GLE. After each team member finished they deliberated on their alignment for each dimension. Each teacher described what they noted for alignment and explained the reasons for making these decisions. When judgments reached consistency, individuals rated independently, completing ratings for the scoring protocol against one grade band of extended grade level expectations. Teams completed one grade band before moving on to the next.

This report contains one section for each subject area. The analyses are described within each section.

ALIGNMENT RESULTS IN READING

Alignment of ExGLEs to PSGLEs

Raters identified the particular PSGLE that linked to the ExGLE under review and rated the link as insufficient, adequate, or well aligned, to evaluate the link between the ExGLEs in Reading and the PSGLEs. Across the elementary grades, 22% of ExGLEs were rated as “well aligned” to the identified, corresponding PSGLE. More “well aligned” ratings were seen in the upper elementary grades than lower elementary grades. In grades 3/4, 45% were either adequately or well aligned, while in grades 5/6, 67% were at least adequately aligned. Overall, there were relatively few ExGLEs for which a PSGLE could not be identified (15% for Grades 3/4, 10% for Grades 5/6). The proportion of ExGLEs rated as “well aligned” in grades 3/4 was 10% and the proportion in grades 5/6 was 33%.

Table 2. Alignment of ExGLEs to PSGLEs – Elementary Reading

ExGLE	PSGLE		ExGLE	PSGLE	
	Grades 3/4	3 rd grade		4 th grade	Grades 5/6
1.1-1 / 2.1-1	○	○	2.1-1	×	○
1.1-2 / 2.1-2	○	○	2.1-2	○	○
1.1-3 / 2.1-3	●	●	2.2-1	●	●
1.1-4 / 2.1-4	×	×	2.2-2	●	●
1.1-5 / 2.1-5	×	×	2.2-3	×	×
1.2 / 2.2-1	●	●	2.3-1	○	○
1.2 / 2.2-2	○	○	2.3-2	×	×
1.3 / 2.3-1	○	○	2.3-3	○	○
1.3 / 2.3-2	×	×	2.4-1	●	●
1.4 / 2.4-1	●	●	2.5-1	●	●
1.5 / 2.5-1	●	●	2.5-2	○	○
1.6 / 2.6-1	●	×	2.6-1	●	●
1.7 / 2.7-1	●	×	2.7-1	●	●
1.7 / 2.7-2	●	×	2.7-2	●	●
1.8 / 2.8-1	●	●	2.8-1	●	●
1.9 / 2.9-1	●	●	2.8-2	●	●
1.9 / 2.9-2	●	●	2.9-1	●	●
1.10 / 2.10-1	×	●	2.9-2	●	●
1.10 / 2.10-2	●	●	2.10-1	●	●
1.11 / 2.11-1	○	●	2.10-2	●	●
			2.11-1	●	○
Percent ●	10	5	Percent ●	38	33
Percent ●	45	45	Percent ●	29	29
Percent ○	25	20	Percent ○	19	29
Percent ×	20	30	Percent ×	14	10

○= Insufficient, ●=Adequate, ●=Well aligned, × = No judgment could be made

Across the middle/secondary grades, 70% of ExGLEs were rated as “well aligned” to the identified, corresponding PSGLE. More “well aligned” ratings were seen in the secondary grades (9/10) than the middle school grades (7/8). In grades 7/8, 95% were either adequately or well aligned, while in grades 9/10, 94% were at least adequately aligned. Overall, there were no middle/secondary ExGLEs for which a PSGLE could not be identified. The proportion of ExGLEs rated as “well aligned” in grades 7/8 was 63% and the proportion in grades 9/10 was 78%.

Table 3. Alignment of ExGLEs to PSGLEs – Middle/Secondary Reading

ExGLE	PSGLE		ExGLE	PSGLE	
	Grades 7/8	7 th grade		8 th grade	Grades 9/10
3.1-1	●	○	4.1-1	●	●
3.1-2	●	●	3.2-1	●	●
3.1-3	●	●	3.2-2	●	●
3.2-1	●	●	4.2-1	●	●
3.2-2	●	●	4.3-1	●	●
3.3-1	●	●	4.3-2	●	●
3.4-1	●	●	4.4-1	●	●
3.4-2	●	●	4.5-1	●	●
3.5-1	●	●	4.5-2	●	●
3.6-1	●	●	4.6-1	●	●
3.6-2	●	●	4.6-2	●	●
3.7-1	●	●	4.6-3	●	●
3.7-2	●	●	4.7-1	●	●
3.7-3	●	●	4.7-2	○	○
3.8-1	●	●	4.8-1	●	●
3.8-2	●	●	4.8-2	●	●
3.9-1	●	●	4.9-1	●	●
3.9-2	●	●	4.9-2	●	●
3.10-1	●	●			
Percent ●	63	53	Percent ●	72	72
Percent ●	37	42	Percent ●	22	22
Percent ○	0	5	Percent ○	5	5
Percent ×	0		Percent ×	0	0

○= Insufficient, ●=Adequate, ●=Well aligned, × = No judgment could be made

Alignment of Extended Reading AA to ExGLEs

Degree of Link. The analyses described in this section of the report address the alignment of the reading alternate assessment (AA) tasks/items (referred to as items in the remainder of the report) to the ExGLEs. Alaska did not specify in the test blueprint for reading their intended alignment of the assessment to the ExGLEs. The degree of alignment here is determined by the link between reading items and the full set of ExGLEs. All AA items from the 2007 assessment were rated four times, once for each grade pair: 3/4, 5/6, 7/8, and 9/10. Reading contains 93 items. Raters were directed to identify the primary link between each assessment item and an ExGLE. These links are called “Fine” grain links and the counts appear in Table 4 in the column headed “Fine.” When a fine grain link could not be identified raters noted this by marking an X. Some Xs were resolved during the consensus process. Any Xs remaining on the third day were reexamined by panel of content experts to identify a link to a content standard. These links were referred to as coarse grain links and are presented in the column headed “Coarse.” In some cases a link could not be found at the fine or coarse grain levels but raters identified a link between the AA item and an ExGLE performance standard from another grade band. These links are reported in the column headed “Other grade band.”

Table 4. Content Alignment of AA items to Elementary ExGLEs (93 items)

	Grades 3/4			Grades 5/6		
	Fine	Coarse	Other grade band	Fine	Coarse	Other grade band
Fluency	13			24		
Word Identification	48	14		13	2	36
Forming a General Understanding	5	2		3		2
Analysis of Content and Structure	1	4	6	7		1
Total	67	20	6	47	2	39

When evaluated against ExGLEs at grades 3/4, 67 of 93 (72%) of reading items had a primary match identified at the fine grain (ExGLE) level. All of the remaining AA items could be matched at a coarser grain (content standard) level, although 6 items were identified with a content standard at a different grade level.

When evaluated against ExGLEs at grades 5/6, 47 of 93 (51%) of reading items had a primary match identified at the fine grain (ExGLE) level. Five of the remaining reading items could not be matched even at a coarser grain (content standard) level: Task 12 Level 1, items 4, 5, and 6 plus Task 12 Level 2, items 5 and 6. There were 39 items (42%) that were matched to another grade band, either a higher or lower grade.

Table 5. Primary Content Alignment of AA items to Middle & Secondary ExGLEs (93 items)

	Grades 7/8			Grades 9/10		
	Fine	Coarse	Other grade band	Fine	Coarse	Other grade band
Fluency	6			3		
Word Identification	10		59	0		69
Forming a General Understanding	0			0		
Analysis of Content and Structure	18			18		
Total	34	0	59	21	0	69

When evaluated against ExGLEs at grades 7/8, 34 of 93 reading items (37%) had a primary match identified at the fine grain (ExGLE) level. None of the remaining reading items could be matched at the coarse grain (content standard) level. There were 59 items (63%) that were matched to another grade pair, either above or below grade 7/8.

At grade pair 9/10, 21 of 93 of reading items (23%) had a primary match identified at the fine grain (ExGLE) level. None of the remaining reading items could be matched, even at a coarse grain (content standard) level, within the 9/10 grade pair ExGLE. There were 69 items (74%) that were matched to another grade band in a lower grade pair. In grade pair 9/10: 3 items could not be aligned even at the coarse grain level, Task 11 Read Passages—timed passages (Simple, Level 1, and Level 2).

Webb Alignment: The remaining Webb statistics are based on the items aligned at the fine grain level within each grade band.

Categorical concurrence is the consistency of categories of content in both the content standards and the assessments. The criterion of categorical concurrence between standards and assessments is met if the same or consistent categories of content appear in both documents (Webb, 1999). Categorical concurrence can be calculated by counting the number of *standards* in which an alternate assessment item appears, then calculating the percentage by dividing the total number of standards with items by the total number of standards. According to Webb (1997), however, to produce an acceptable level of reliability for assessment scores, there should be at least six items per standard. In general assessments, to meet this criterion, categorical concurrences values are presented as the percentage of content standards with at least six items aligned. Both indices are reported in Table 6.

Table 6. Categorical Concurrence Summary, Reading

Grade	% of Standards w/ at least one item	Standards not represented*	% of standards w/ six or more items	Standards that are under-represented*
3/4	36%	1.4, 1.5, 1.6, 1.7, 1.9, 1.10, 1.11	18%	1.2, 1.8
5/6	45%	2.4, 2.5, 2.6, 2.7, 2.10, 2.11	27%	2.2, 2.9
7/8	50%	3.3, 3.4, 3.5, 3.6, 3.10	30%	3.8, 3.9
9/10	20%	4.1, 4.2, 4.3, 4.4, 4.5, 4.7, 4.8, 4.9	10%	3.2

*PSGLE codes

Depth of Knowledge (DOK) examines the consistency between the cognitive demands of the standards and cognitive demands of corresponding assessment items. Completely aligned standards and assessments require an assessment system designed to measure in some way the full range of expected knowledge within each specified content standard. The percent of AA items below, at, or above the standards to which they are being compared are reported in the table below. In grades 3/4 and 9/10, none of the assessment items had DOK ratings below those of the corresponding ExGLE.

Table 7. Percentage of AA items below, at, and above the ExGLEs, Reading

Grade	# items evaluated	% Below	% At	% Above	% At or Above
3/4	67	0.0	76.1	23.9	100
5/6	47	34.0	44.7	21.3	66
7/8	34	29.4	58.8	11.8	71
9/10	21	0.0	57.1	42.9	100

Balance of representation is examined by reviewing the emphasis the assessment gives to different objectives within a content standard. The content standards and assessments give comparable emphasis to what students are expected to know, what they should be able to do, and in what contexts they are expected to demonstrate their proficiency. The balance of representation criterion is used to indicate the extent to which items are evenly distributed across content standards. The formula used to compute the balance of representation index is

$$Balance = 1 - \left(\sum_{i=1}^k \left| \frac{1}{O} - \frac{I_k}{H} \right| \right) / 2,$$

where O is the total number of objectives hit for the standard, I_k is the number of items hit corresponding to objective k , and H is the total number of items hit for the content standard. Index values close to zero would indicate unbalanced item representation across the content standards and values close to 1.0 indicate balanced representation. In general assessments, index values between .6 and .7 indicate balance of representation criterion has only been weakly met and values greater than .7 indicate that the criterion has been met. Balance of representation in reading ranged from .19 in grades 5/6 to .58 in grades 9/10.

Table 8. Balance of Representation of AA Items vs. Grade Band ExGLEs

	Fluency	Word ID	Form Underst.	Analysis	Sum	Overall Balance [1-sum/2]
3/4	.24	.39	.50	.0	1.13	.43
5/6	.24	.67	.33	.39	1.61	.19
7/8	0	.60	0	.45	1.05	.48
9/10	.50	.0	.0	.34	.84	.58

The *range-of-knowledge correspondence* criterion examines the alignment of assessment items to the multiple objectives within the content standards. Range-of-knowledge correspondence is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities. The range-of-knowledge value is the percentage of objectives that have at least one hit (i.e., an assessment item aligned to that objective). In general assessments, an acceptable value for range-of-knowledge correspondence is identified as at least 50% of objectives having at least one aligned assessment item. Range of knowledge correspondence for reading ranged from 22% in grades 9/10 to 47% in grades 7/8.

Table 8-1. Range-of-Knowledge Correspondence, Reading Items vs. Grade Band ExGLEs

	Grades 3/4	Grades 5/6	Grades 7/8	Grades 9/10
# of ExGLEs	20	21	19	18
# w/ at least one match	7	8	9	4
% range of knowledge correspond	35%	38%	47%	22%

ALIGNMENT RESULTS IN WRITING

Alignment of ExGLEs to PSGLEs

Raters identified the particular PSGLE that linked to the ExGLE under review and rated the link as insufficient, adequate, or well aligned, to evaluate the link between the ExGLEs in Writing and the PSGLEs. Table 9 shows the results for the elementary ExGLEs, grade bands 3/4 and 5/6. Two of the ExGLEs (29%) were linked to PSGLEs in grade 3/4 and 2 ExGLEs (40%) were minimally linked in grade 5/6. No ExGLEs were rated as well aligned to PSGLEs in the elementary grades.

Table 9. Alignment of ExGLEs to PSGLEs – Elementary Writing

ExGLE	PSGLE		ExGLE	PSGLE	
	3 rd grade	4 th grade		Grades 5/6	5 th grade
1.1-1 / 2.1-1	×	×	2.1-1	×	×
1.1-2 / 2.1-2	×	×	2.1-2	×	×
1.1-3 / 2.1-3	○	×	2.2-1	●	●
1.2-1 / 2.2-1	×	×	2.3-1	●	○
1.3 / 2.3-1	○	○	2.4-1	○	○
1.3 / 2.3-2	●	●			
1.4-1	●	●			
Percent ●	0	0	Percent ●	0	0
Percent ●	29	29	Percent ●	40	20
Percent ○	29	14	Percent ○	20	40
Percent ×	43	57	Percent ×	40	40

○= Insufficient, ●=Adequate, ●=Well aligned, × = No judgment could be made

Table XX shows the mid/secondary grade band ExGLEs. All ExGLEs in grade bands 7/8 and 9/10 were adequately linked to PSGLEs and one ExGLE in each mid/secondary grade band (7/8 and 9/10) was rated well aligned. Overall, linking between the elementary ExGLEs and PSGLEs was insufficient and linking in the mid/secondary grade bands was rated adequate or better.

Table 10. Alignment of ExGLEs to PSGLEs – Middle/Secondary Writing

ExGLE	PSGLE		ExGLE	PSGLE	
	7 th grade	8 th grade		Grades 9/10	9 th grade
3.1-1	●	●	4.1-1	●	●
3.2-1	●	●	4.2-1	●	●
3.3-1	●	●	4.3-1	●	●
3.4-1	●	●	4.4-1	●	●
3.4-2	●	●	4.4-2	●	●
			4.4-3	●	●
Percent ●	20	0	Percent ●	17	17
Percent ●	80	100	Percent ●	83	83
Percent ○	0	0	Percent ○	0	0
Percent ×	0	0	Percent ×	0	0

○= Insufficient, ●=Adequate, ●=Well aligned, × = No judgment could be made

Note that there are no state-assessed ExGLEs for two of the five writing strands in the performance (*content*) standards, the standards for citing sources and using resources. Although, neither standard is to be assessed on Alaska’s alternate assessment, publishing ExGLEs for these strands may guide local assessments and may also be useful for teachers in planning instruction.

Alignment of AA to ExGLEs

The analyses described in this section of the report address the alignment of the writing alternate assessment (AA) tasks/items (referred to as items in the remainder of the section) to the ExGLEs. Alaska has not yet specified in the test blueprint for writing, their intended content emphasis within the writing AA related to the performance (*content*) standards in the ExGLEs. The degree of alignment in this report is determined by the link between the writing AA items and all ExGLEs except for those in the two strands intended for local assessment (revising and citing sources). All AA items from the 2007 writing assessment were rated four times, once for each of the ExGLE grade bands: 3/4, 5/6, 7/8, and 9/10. Results reported here are based on 54 items. Some items were functions of others (# correct, percent correct) so only the base item, not the derivative, was rated.

Degree of Link. Raters were directed to identify the primary link between each assessment item and an ExGLE. These links are called “Fine” grain links and the counts appear in Table 11 under the column headed “Fine.” When a fine grain link could not be identified raters noted this by marking an X. Some Xs were resolved during the consensus process. Any Xs remaining on the third day were reexamined by panel content experts to identify a link to a content standard. These links were referred to as coarse grain links and are presented in the column headed “Coarse.” In some cases a link could not be found at the fine or coarse grain levels but raters identified a link between the AA item and an ExGLE performance (*content*) standard from another grade band. These links are reported in the column headed “Other grade band.”

Table 11 shows the degree of link for the elementary grade bands. When evaluated against the ExGLEs at grade band 3/4, 48 of the 54 items (88%) had a primary match identified at the fine grain level. Six items were not aligned at the coarse grain and were not aligned to another grade

band. These items included Task 7, Sentence Mechanics, Item 4; Task 8, Item 1; Task 9 Write Story (Pictures), Items 1 and 6; and Task 10, Write Story, items 1 and 4. When evaluated against ExGLEs at grades 5/6, 54 of 54 items (100%) had a primary match identified at the fine grain (ExGLE) level.

Table 11. Content Alignment of AA Writing Items to Elementary ExGLEs (54 items)

	Grades 3/4			Grades 5/6		
	Fine	Coarse	Other grade band	Fine	Coarse	Other grade band
Writing using a variety of forms	36			23		
Structures and conventions	1			25		
Revising	11			6		
Total	48	0	0	54	N/A	N/A

Table 12 shows the degree of link for the mid/secondary grade bands. When evaluated against the ExGLEs at grade band 7/8, 47 of the 54 writing items (87%) had a primary match at the fine grain level. Five of the remaining items were matched at the coarse grain level. There were two items that were not aligned even at the coarse grain level. These were Task 9, Write a Story (Pictures), Item 1 and Task 10, Write a Story, Item 1.

At grade band 9/10, 41 of 54 items (76%) had a primary match at the fine grain level and 10 of 54 items (19%) had a primary match at the coarse grain level. There were three items that were not aligned even at the coarse grain. These were Task 8, Write a Sentence, Item 1; Task 9, Write a Story (Pictures), Item 1; and Task 10, Write a Story, Item 1.

Table 12. Content Alignment of AA items to Middle & Secondary ExGLEs (54 items)

	Grades 7/8			Grades 9/10		
	Fine	Coarse	Other grade band	Fine	Coarse	Other grade band
Writing using a variety of forms	4			6		
Structures and conventions	43	5		27	10	
Revising	0			8		
Total	47	5	0	41	10	0

Note: Three of the writing AA tasks ask the student to create a piece of writing. In Task 8, the student writes a sentence about a given prompt, in Task 9 the student writes a story with picture prompts, and in Task 10, the student writes a story without picture prompts. The assessor records the total number of words written for each task. This item, total words, triggered much discussion among the writing experts. During the coarse grain review on the third day, all writing

content experts discussed the item. They indicated that as an assessment item, it “needs to include a qualitative piece, not just number, but also a story is told.” The metric for these items breaks scoring into a quantitative component, total words, and a qualitative component, ideas and organization. Content experts were inconsistent in their judgments, rating the total word item as an isolated element and unsure about identifying the best match.

Webb Alignment: The remaining Webb statistics are based on the items aligned at the fine grain level within each grade band.

Categorical concurrence is the consistency of categories of content in both the content standards and the assessments. The criterion of categorical concurrence between standards and assessment is met if the same or consistent categories of content appear in both documents (Webb, 1999). Categorical concurrence can be calculated by counting the number of *standards* in which an alternate assessment item appears, then calculating the percentage by dividing the total number of standards with items by the total number of standards. According to Webb (1997), however, to produce an acceptable level of reliability for assessment scores, there should be at least six items per standard. In general assessments, to meet this criterion, categorical concurrences values are presented as the percentage of content standards with at least six items aligned. Both indices are reported in Table 13. Note that the locally (L) assessed strands were not included in these calculations.

Table 13. Categorical Concurrence Summary, Writing – BASED ON STANDARDS

Grade	% of Standards w/ at least one item	Standards not represented*	% of standards w/ six or more items	Standards that are under-represented*
3/4	100%	N/A	75%	1.3
5/6	100%	N/A	75%	2.2
7/8	75%	3.4	25%	3.1, 3.2
9/10	100%	N/A	50%	4.1, 4.2

*PSGLE codes

Depth of Knowledge (DOK) examines the consistency between the cognitive demands of the standards and cognitive demands of assessment items. Completely aligned standards and assessments require an assessment system designed to measure in some way the full range of expected knowledge within each specified content standard. The percent of AA items below, at, or above the standards to which they are being compared are reported in Table 14.

Table 14. Percentage of AA items below, at, and above the ExGLEs, Writing

Grade	# items evaluated	% Below	% At	% Above	% At or Above
3/4	48	29	50	21	71
5/6	54	72	15	13	28
7/8	47	38	34	28	62
9/10	41	51	48	0	48

Balance of representation is examined by reviewing the emphasis the assessment gives to different objectives within a content standard. The content standards and assessments give

comparable emphasis to what students are expected to know, what they should be able to do, and in what contexts they are expected to demonstrate their proficiency. The balance of representation criterion is used to indicate the extent to which items are evenly distributed across content standards. The formula used to compute the balance of representation index is

$$Balance = 1 - \left(\sum_{i=1}^k \left| \frac{1}{O} - \frac{I_k}{H} \right| \right) / 2,$$

where O is the total number of objectives hit for the standard, I_k is the number of items hit corresponding to objective k , and H is the total number of items hit for the content standard. Index values close to zero would indicate unbalanced item representation across the content standards and values close to 1.0 indicate balanced representation. Index values between .6 and .7 indicate balance of representation criterion has only been weakly met and values greater than .7 indicate that the criterion has been met. Balance of representation indices for Writing ranged from .42 in grades 7/8 to .75 in grades 3/4.

Table 15. Balance of Representation of Writing AA Items vs. Grade Band ExGLEs

	Write Using Variety of Forms	Structures and Conventions	Revise	Sum	Overall Balance [1-sum/2]
3/4	.28	.18	.03	.49	.75
5/6	.43	.26	.09	.79	.60
7/8	.58	.58	0	1.16	.42
9/10	.35	.41	.05	.83	.59

The *range-of-knowledge correspondence* criterion examines the alignment of assessment items to the multiple objectives within the content standards. Range-of-knowledge correspondence is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities. The range-of-knowledge value is the percentage of objectives that have at least one hit (i.e., an assessment item aligned to that objective). In general assessments, an acceptable value for range-of-knowledge correspondence is identified as at least 50% of objectives having at least one aligned assessment item. This criterion was met for all four subjects in writing (see Table XX).

Table 15-2. Range-of-Knowledge Correspondence, Writing Items vs. Grade Band ExGLEs

	Grades 3/4	Grades 5/6	Grades 7/8	Grades 9/10
# of ExGLEs	7	5	5	6
# w/ at least one match	5	5	3	4
% range of knowledge correspond	71%	100%	60%	67%

ALIGNMENT RESULTS IN MATHEMATICS

ALIGNMENT RESULTS IN MATHEMATICS

Alignment of ExGLEs to PSGLEs

Across the elementary grade bands, 10% of ExGLEs were rated as “well aligned” to the identified, corresponding PSGLEs, and for the mid/secondary grade bands 14% were well aligned. In grade band 3-4, 57% and in band 5/6, 48% of ExGLEs were either adequately or well aligned to the corresponding PSGLEs. In the mid/secondary area, grade band 7/8 had 81% alignment between the ExGLEs and the PSGLEs and ExGLEs in grade band 9/10 were rated as being 39% aligned with corresponding PSGLEs. Overall, across all grade bands, 44% of ExGLEs were insufficiently or not aligned to a corresponding PSGLE. Tables 16 and 17 display the grade by grade alignment between the ExGLE grade bands and the corresponding PSGLEs.

Table 16. Alignment of ExGLEs to PSGLEs – Elementary Math

ExGLE	PSGLE		ExGLE	PSGLE	
	3 rd grade	4 th grade		Grades 5/6	5 th grade
N-1	○	○	N-1	○	○
N-2	○	○	N-2	◐	○
N-3	○	○	N-3	○	○
N-4	○	○	N-4	○	○
MEA-1	×	×	N-5	○	○
E&C-1	×	×	N-6	◐	◐
E&C-2	◐	◐	N-7	◐	◐
F&R-1	×	×	MEA-1	×	×
F&R-2	◐	◐	MEA-2	×	×
F&R-3	×	×	E&C-1	×	×
G-1	◐	×	E&C-2	◐	◐
G-2	●	●	F&R-1	○	○
G-3	◐	◐	F&R-2	×	×
G-4	◐	×	F&R-3	×	×
G-5	◐	◐	F&R-4	×	×
S&P-1	×	×	G-1	×	×
S&P-2	◐	◐	G-2	◐	◐
S&P-3	◐	○	G-3	◐	◐
PS-1	●	●	G-4	×	×
PS-2	●	◐	G-5	◐	◐
PS-3	●	◐	S&P-1	×	×
			S&P-2	◐	◐
			S&P-3	◐	×
			PS-1	●	●
			PS-2	×	×

ExGLE	PSGLE		ExGLE	PSGLE		
	Grades 3/4	3 rd grade		4 th grade	Grades 5/6	5 th grade
Percent ●	19	10	Percent ●	7	7	
Percent ◐	38	33	Percent ◐	41	34	
Percent ○	19	24	Percent ○	17	21	
Percent ×	24	33	Percent ×	34	38	

○= Insufficient, ◐=Adequate, ●=Well aligned, × = No judgment could be made

Table 17. Alignment of ExGLEs to PSGLEs – Middle/Secondary Math

ExGLE	PSGLE		ExGLE	PSGLE	
	Grades 7/8	7 th grade		8 th grade	Grades 9/10
N-1	×	●	N-1	×	×
N-2	◐	◐	N-2	×	×
N-3	●	○	N-3	×	×
N-4	◐	●	N-4	●	◐
N-5	●	◐	N-5	×	○
N-6	◐	●	MEA-1	×	×
N-7	×	×	MEA-2	×	×
MEA-1	◐	×	E&C-1	×	×
MEA-2	×	×	E&C-2	◐	○
MEA-3	×	×	E&C-3	×	×
E&C-1	●	◐	F&R-1	◐	◐
E&C-2	×	◐	F&R-2	◐	◐
F&R-1	●	●	F&R-3	×	×
F&R-2	×	×	F&R-4	×	×
F&R-3	○	◐	F&R-5	×	×
G-1	◐	◐	G-1	×	×
G-2	◐	◐	G-2	○	○
G-3	◐	◐	G-3	◐	◐
G-4	○	○	G-4	×	×
G-5	◐	○	G-5	×	×
G-6	◐	◐	G-6	○	○
S&P-1	○	○	G-7	×	×
S&P-2	◐	◐	G-8	○	○
S&P-3	◐	◐	S&P-1	◐	◐
PS-1	◐	×	S&P-2	◐	◐
PS-2	◐	◐	S&P-3	◐	○
PS-3	◐	◐	PS-1	×	×
PS-4	◐	◐	PS-2	○	○

ExGLE	PSGLE		ExGLE	PSGLE	
	7 th grade	8 th grade		Grades 9/10	9 th grade
PS-5	●	●	PS-3	×	×
PS-6	●	●	PS-4	●	●
PS-7	●	●	PS-5	●	●
			PS-6	●	●
			PS-7	●	●
Percent ●	19	19	Percent ●	6	0
Percent ●	52	48	Percent ●	30	30
Percent ○	10	13	Percent ○	12	21
Percent ×	19	19	Percent ×	52	48

○= Insufficient, ●=Adequate, ●=Well aligned, × = No judgment could be made

Alignment of AA to ExGLEs

Degree of Link. The analyses described in this section of the report address the alignment of the mathematics alternate assessment (AA) tasks/items (referred to as items in the remainder of the section) to the ExGLEs. Alaska has not specified in the test blueprint for the mathematics alternate assessment their intended alignment of the assessment to the ExGLEs. The degree of alignment reported here is determined by the link between mathematics items and all ExGLEs designated for the state-level assessment. All mathematics AA items from the 2007 assessment were rated four times, once for each grade pair: 3/4, 5/6, 7/8, and 9/10. Mathematics contains 104 items. Raters were directed to identify the primary link between each assessment item and an ExGLE. These links are called “Fine” grain links and the counts for the elementary grade bands appear in Table 18 in the column headed “Fine.” When a fine grain link could not be identified raters noted this by marking an X. Some Xs were resolved during the consensus process. Any Xs remaining on the third day were reexamined by panel content experts to identify a link to a content standard. These links were referred to as coarse grain links and are presented in the column headed “Coarse.” There were a large number of items that could not be aligned at the fine grain size in the area of measurement. These were grouped together in the coarse grain category.

Table 18. Content Alignment of AA items to Elementary ExGLEs (104 items)

	<u>Grades 3/4</u>		<u>Grades 5/6</u>	
	Fine	Coarse	Fine	Coarse
Numeration	18		32	
Measurement		66	8	59
Estimation/Computation	1		4	
Functions & Relationships				
Geometry	12		12	
Statistics & Probability				
Problem Solving			2	
Total	31	66	58	59

Table 19 displays the ratings for mid/secondary grade bands.

Table 19. Content Alignment of AA items to Middle & Secondary ExGLEs (104 items)

	<u>Grades 7/8</u>		<u>Grades 9/10</u>	
	Fine	Coarse	Fine	Coarse
Numeration	23		7	
Measurement	20	39	13	50
Estimation/Computation				
Functions & Relationships			1	
Geometry	12			
Statistics & Probability				
Problem Solving				
Total	55	39	21	50

In some cases a link could not be found at the fine or coarse grain levels. These items are listed in Table 20.

Table 20. Items that could not be aligned even at the coarse grain level

Grade Band	# of Xs	Math Tasks	Items
3/4	9	9 Number Line	1
		10 Tell Time	1, 2, 3
		16 Fractions	1, 2, 3, 4, 5
5/6	12	7 Identify Money	5, 6, 7, 8, 9
		9 Number Line	1
		13 Calendar	3, 4, 5
		16 Fractions	1, 2, 3
7/8	9	13 Calendar	1, 2, 3, 4, 5
		16 Fractions	1, 3, 5
		18 Place Value	3
9/10	31	5 Discriminate Differences	1, 2, 3, 4
		7 Identify Money	1, 2, 3, 4, 5, 6, 7, 8
		9 Number Line	2, 4, 5
		13 Calendar	1, 2, 3, 4, 5
		14 Count/Take Away	1, 2, 3, 4, 5, 6, 7
		20 Addition	1-20
		21 Subtraction	1-20

Webb Alignment: The remaining Webb statistics are based on the items aligned at the fine grain level within each grade band.

Categorical concurrence is the consistency of categories of content in both the content standards and the assessments. The criterion of categorical concurrence between standards and assessments is met if the same or consistent categories of content appear in both documents (Webb, 1999). Categorical concurrence can be calculated by counting the number of *standards* in which an alternate assessment item appears, then calculating the percentage by dividing the total number of standards with items by the total number of standards. According to Webb (1997), however, to produce an acceptable level of reliability for assessment scores, there should be at least six items per standard. In general assessments, to meet this criterion, categorical concurrences values are presented as the percentage of content standards with at least six items aligned. Both indices are reported in Table 21. Note that although some standards are not designated to be assessed on the state assessment. There are, however, a few mathematics AA items that tap the local (L) ExGLEs. The standards flagged to be locally assessed are not included in the base analysis but are reported in the final “under-represented” column.

Table 21. Categorical Concurrence Summary, Math

Grade	% of Standards w/ at least one item	Standards not represented*	% of standards w/ six or more items	Standards that are under-represented^
3/4	67%	N1, FR2, PS2-4	33%	N3, (EC1), G2
5/6	69%	EC2, FR1, G3-4, SP2, PS4	13%	N1, N3-4, MEA1-2, (EC1), G2, (PS1), PS4
7/8	52%	EC2, FR1-3, G2, G4-5, SP2, PS5-7	14%	N1, N3-5, N7, MEA2-3, (G6)
9/10	25%	MEA2, EC1-3, FR1, FR3, G1-2, G4-5, SP1-2, PS5-7	5%	N1-3, FR1

*List includes only the ExGLEs intended to be measured by the state

^ Parentheses indicate (L) standard that was measured anyway

Depth of Knowledge (DOK) examines the consistency between the cognitive demands of the standards and cognitive demands of assessment items. Completely aligned standards and assessments require an assessment system designed to measure in some way the full range of expected knowledge within each specified content standard. The percent of AA items below, at, or above the standards to which they are being compared are reported in Table 22.

Table 22. Percentage of Extended math items below, at, and above the ExGLEs

Grade	# items evaluated	% Below	% At	% Above	% At or Above
3/4	31	29	58	13	71
5/6	58	33	55	12	67
7/8	55	16	49	35	84
9/10	21	48	24	29	53

Balance of representation is examined by reviewing the emphasis the assessment gives to different objectives within a content standard. The content standards and assessments give comparable emphasis to what students are expected to know, what they should be able to do, and in what contexts they are expected to demonstrate their proficiency. The balance of representation criterion is used to indicate the extent to which items are evenly distributed across content standards. The formula used to compute the balance of representation index is

$$Balance = 1 - \left(\sum_{i=1}^k \left| \frac{1}{O} - \frac{I_k}{H} \right| \right) / 2,$$

where O is the total number of objectives hit for the standard, I_k is the number of items hit corresponding to objective k , and H is the total number of items hit for the content standard. Index values close to zero would indicate unbalanced item representation across the content standards and values close to 1.0 indicate balance representation. In general assessments, index values between .6 and .7 indicate balance of representation criterion has only been weakly met, and values greater than .7 indicate that the criterion has been met. Math items met this criterion when evaluated against all grade bands except 9/10. Note that these statistics do not reflect the large number of measurement items identified at the coarser grain.

Table 23-1. Balance of Representation of AA Items vs. Grade Band ExGLEs

	Num	Mea	EC	FR	G	SP	PS	Sum	Overall Balance [1-sum/2]
3/4	.28		.13		.13			.55	.73
5/6	.31	.04	.02		.07		.14	.60	.70
7/8	.23	.17			.07			.47	.76
9/10	.34	.42		.15				.91	.54

The *range-of-knowledge correspondence* criterion examines the alignment of assessment items to the multiple objectives within the content standards. Range-of-knowledge correspondence is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities. The range-of-knowledge value is the percentage of objectives that have at least one hit (i.e., an assessment item aligned to that objective). In general assessments, an acceptable value for range-of-knowledge correspondence is identified as at least 50% of objectives having at least one aligned assessment item. This indicator ranged from 25% in grades 9/10 to 69% in grades 5/6.

Table 23-2. Range of Knowledge Correspondence, Extended Math Items vs. State-Assessed ExGLEs

	Grades 3/4	Grades 5/6	Grades 7/8	Grades 9/10
# of ExGLEs*	9	16	21	20
# w/ at least one match	6	11	11	5
% range of knowledge correspond	67%	69%	52%	25%

*Excludes locally assessed

ALIGNMENT RESULTS IN SCIENCE

Alignment of ExGLEs to PSGLEs

Raters identified the particular PSGLE that linked to the ExGLE under review and rated the link as insufficient, adequate, or well aligned, to evaluate the link between the ExGLEs in Science and the PSGLEs. Table 24 shows the results for the elementary ExGLEs, grade bands 3/4 and 5/6. Table 25 shows the mid/secondary grade band ExGLEs. All ExGLEs in grade bands 7/8 and 9/10 were adequately linked to PSGLEs and one ExGLE in each mid/secondary grade band (7/8 and 9/10) was rated well aligned. Both the elementary mid/secondary ExGLEs were linked to the PSGLEs in science ranging from 62% (9/10) to 71% (3/4 and 7/8) of ExGLEs with at least one “adequate” judgment.

Table 24. Alignment of ExGLEs to PSGLEs – Elementary Science

ExGLE	PSGLE		ExGLE	PSGLE	
	Grades 3/4	3 rd grade 4 th grade		Grades 5/6	5 th grade 6 th grade
SA1.1	●	●	SA1.1	●	●
SA1.2	●	●	SA1.2	●	●
SA3.1	●	●	SA3.1	●	●
SB1.1	○	●	SB1.1	●	○
SB2.1	×	×	SB2.1	●	●
SB3.1	●	●	SB3.1	●	●
SB4.1	●	●	SB4.1	○	×
SC1.1	●	●	SC1.1	●	●
SC2.1	●	●	SC2.1	×	×
SC3.1	●	●	SC3.1	●	×
SD1.1	×	×	SD1.1	●	●
SD1.2	○	×	SD1.2	●	×
SD2.1	●	×	SD2.1	○	●
SD3.1	×	×	SD3.1	×	×
SD4.1	●	●	SD4.1	●	×
SE1.1	●	×	SE1.1	●	●
SE2.1	●	●	SE2.1	○	○
SE3.1	○	×	SE3.1	○	○
SF1.1	●	●	SF1.1	●	●
SG1.1	●	●	SG1.1	×	×
SG2.1	○	×	SG2.1	×	×
Percent ●	33	19	Percent ●	5	5
Percent ●	33	43	Percent ●	57	43
Percent ○	19	0	Percent ○	19	14
Percent ×	14	38	Percent ×	19	38

○= Insufficient, ●=Adequate, ●=Well aligned, × = No judgment could be made

Table 25. Alignment of ExGLEs to PSGLEs – Middle/Secondary Science

ExGLE	PSGLE		ExGLE	PSGLE	
	Grades 7/8	7 th grade		8 th grade	Grades 9/10
SA1.1	●	●	SA1.1	●	●
SA1.2	●	●	SA1.2	●	●
SA3.1	●	●	SA3.1	×	×
SB1.1	●	●	SB1.1	●	●
SB2.1	○	○	SB2.1	○	○
SB3.1	●	●	SB3.1	×	×
SB4.1	●	○	SB4.1	●	×
SC1.1	●	●	SC1.1	×	●
SC2.1	×	×	SC2.1	○	●
SC3.1	●	●	SC3.1	○	●
SD1.1	×	×	SD1.1	●	●
SD1.2	●	●	SD1.2	●	●
SD2.1	●	×	SD2.1	●	●
SD3.1	×	×	SD3.1	●	×
SD4.1	×	×	SD4.1	×	×
SE1.1	●	●	SE1.1	●	●
SE2.1	×	×	SE2.1	×	×
SE3.1	●	●	SE3.1	●	○
SF1.1	●	●	SF1.1	○	○
SG1.1	●	●	SG1.1	×	×
SG2.1	●	●	SG2.1	×	×
Percent ●	14	24	Percent ●	29	33
Percent ●	57	38	Percent ●	19	14
Percent ○	5	10	Percent ○	19	14
Percent ×	24	29	Percent ×	33	38

○= Insufficient, ●=Adequate, ●=Well aligned, × = No judgment could be made

Alignment of AA to ExGLEs

The analyses described in this section of the report address the alignment of the science alternate assessment (AA) tasks/items (referred to as items in the remainder of the section) to the ExGLEs for three grade pairs because assessments in science will be given at three grades in 2008, grades 4, 8, and 10. Raters were directed to identify the primary link between each assessment item and an ExGLE. It is important to note that locally assessed strands, including Science as Inquiry and Process as well as Cultural, Social, and Personal Perspectives and Science were excluded from this analysis. The links and the item counts appear in Table XX. Results are based on 55 items.

Table 26. Primary Content Alignment of Extended Science Items to ExGLEs (55 items)

	3/4	7/8	9/10
Physical Science	16	19	19
Life Science	22	21	21
Earth Science	10	8	10
Science & Technology	5	5	5
History & Nature of Science	2	0	0
Total	55	53	55

Task 10, Change in Lithosphere, Items 3 and 5 could not be clearly aligned at the fine grain to ExGLEs in grades 7/8.

Webb Alignment: The remaining Webb statistics are based on the items aligned at the fine grain level within each grade band.

Categorical concurrence is the consistency of categories of content in both the content standards and the assessments. The criterion of categorical concurrence between standards and assessments is met if the same or consistent categories of content appear in both documents (Webb, 1999). To produce an acceptable level of reliability for assessment scores, Webb (1997) recommends at least six items per standard. Categorical concurrences values are the percentage of content standards with at least six items aligned. Inquiry & Culture strands are omitted from the total because those are locally assessed. History & Nature is included because it is locally assessed at only some grade bands.

Table 27. Categorical Concurrence Summary, Science – based on state-assessed standards (N=17)

Grade	% of Standards w/ at least one item	Standards not represented*	% of standards w/ six or more items	Standards that are under-represented*
3/4	65	SD1.1, 2.1, 4.1, SE1.1, 2.1, SG2.1	18	SB1.1, 2.1, 3.1 SC1.1, 3.1 SD1.2, SE3.1, SG1.1
7/8	65	SD2.1, 4.1, SE1.1, 2.1, SG1.1, 2.1	24	SB1.1, 2.1, SC1.1, 3.1 SD1.1, 1.2, 3.1 SE3.1
9/10	65	SD1.1, 2.1, 4.1, SE2.1, SG1.1, 2.1	29	SB1.1, 2.1, SC1.1, SD1.1, SE1.1, 3.1

*PSGLE codes

Depth of Knowledge (DOK) examines the consistency between the cognitive demands of the standards and cognitive demands of assessment items. Completely aligned standards and assessments requires an assessment system designed to measure in some way the full range of expected knowledge within each specified content standard. The percent of AA items below, at, or above the standards to which they are being compared are reported in the table below.

Table 28. Percentage of Extended Science Items below, at, and above the ExGLEs

Grade	# items evaluated	% Below	% At	% Above	% At or Above
3/4	55	42	22	36	58
7/8	53	34	13	53	66
9/10	55	56	26	18	44

Balance of representation is examined by reviewing the emphasis the assessment gives to different objectives within a content standard. The content standards and assessments give comparable emphasis to what students are expected to know, what they should be able to do, and in what contexts they are expected to demonstrate their proficiency. The balance of representation criterion is used to indicate the extent to which items are evenly distributed across content standards. The formula used to compute the balance of representation index is

$$\text{Balance} = 1 - \left(\sum_{i=1}^k \left| \frac{1}{O} - \frac{I_k}{H} \right| \right) / 2,$$

where O is the total number of objectives hit for the standard, I_k is the number of items hit corresponding to objective k , and H is the total number of items hit for the content standard. Index values close to zero would indicate unbalanced item representation across the content standards and values close to 1.0 indicate balanced representation. Index values between .6 and .7 indicate balance of representation criterion has only been weakly met and values greater than .7 indicate that the criterion has been met. Balance of representation values were fairly consistent across the grade bands.

Table 29. Balance of Representation of Extended Science Items vs. Grade Band ExGLEs

	Physical	Life	Earth	Tech	History	Sum	Overall Balance [1-sum/2]
3/4	.12	.37	.10		.05	.65	.67
7/8	.09	.22	.12			.45	.78
9/10	.10	.25	.24	.04		.64	.68

The **range-of-knowledge correspondence** criterion examines the alignment of assessment items to the multiple objectives within the content standards. Range-of-knowledge correspondence is used to judge whether a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities. The range-of-knowledge value is the percentage of objectives that have at least one hit (i.e., an assessment item aligned to that objective). An acceptable value for range-of-knowledge correspondence is identified as at least 50% of objectives having at least

one aligned assessment item. Range of knowledge correspondence was 65% across all three grade bands.

Table 30. Range-of-Knowledge Correspondence, Extended Science Items vs. Grade Band ExGLEs

	Grades 3/4	Grades 7/8	Grades 9/10
# of ExGLEs	17	17	17
# w/ at least one match	11	11	11
% range of knowledge correspond	65%	65%	65%

REFERENCES

- Flowers, C., Browder, D., & Ahlgrim-Delzell, L. (2006). An analysis of three states' alignment between language arts and mathematics standards and alternate assessments. *Exceptional Children*, 72(2), 201-215.
- Karvonen, M., Flowers, C. P., Wakeman, S. L., & Browder, D. M. (2006, April). *Measuring the enacted curriculum for students with significant cognitive disabilities*. Paper in an invited symposium for the 2006 conference of the National Council on Measurement in Education, San Francisco, CA.
- La Marca, P. M. (2001). Alignment of standards and assessments as an accountability criterion. *Practical Assessment, Research & Evaluation*, 7(21). Retrieved June 4, 2003, from <http://ericae.net/pare/getvn.asp?v=7&n=21>
- Roach, A. T., Elliott, S. N., Webb, N. L. (2005). Alignment of an alternate assessment with state academic standards: Evidence for the content validity of the Wisconsin alternate assessment. *The Journal of Special Education*, 38(4), 218-231.
- Tindal (2005). *Alignment of alternate assessments using the Webb system*. Washington, DC: Council of Chief State School Officers.
- US Department of Education (2005, August). Alternate achievement standards for students with the most significant cognitive disabilities non-regulatory guidance. Retrieved from the World Wide Web on October 18, 2005 at <http://www.ed.gov/policy/elsec/guid/altguidance.pdf>.
- Webb, N. L. (1997). *Determining alignment of expectations and assessments in mathematics and science education. NISE Brief*. National Center for Improving Science Education, University of Wisconsin–Madison.
- Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education*. Madison, WI: University of Wisconsin-Madison, National Institute for Science Education.
- Webb, N. L. (2002). *Alignment study in language arts, mathematics, science, and social studies of state standards and assessments for four states*. Washington, D.C.: Council of Chief State School Officers.

APPENDIX

Dr. Meagan Karvonen, Western Carolina University

Dr. Karvonen is an assistant professor of educational research at Western Carolina University. She received her Ph.D. in educational research and measurement from the University of South Carolina and teaches graduate level courses in research methodology. Dr. Karvonen began conducting research on alternate assessment systems in 2000, through a federally-funded project to help teachers create access to the general curriculum and construct high-quality portfolio alternate assessments. Since that time she has worked on several other federally-funded projects related to alternate assessments. She currently serves as a research associate for the National Alternate Assessment Center, working with the group at the University of North Carolina at Charlotte on developing and applying methodologies for investigating the alignment of alternate assessments to other elements of the educational system. She has conducted alignment studies on alternate assessments in several states. Dr. Karvonen is a co-author of more than 30 publications, presentations, and other manuscripts on general curriculum access, alternate assessment, and alignment. In 2005 she was a contributing author to the *Tool Kit on Teaching and Assessing Students with Disabilities*, commissioned by the U.S. Department of Education.

Dr. Patricia Almond, Behavioral Research and Teaching, University of Oregon

Dr. Almond is currently a research associate with Behavioral Research and Teaching center at the University of Oregon. She was a co-principal investigator with the Developing Alternate Assessment Technical Adequacy project, with the Assessing Special Education Students collaborative. Between 1996 and 2004 Dr. Almond was director of alternate assessments and testing accommodations for the Oregon Department of Education. She directed the design, development, and implementation of Oregon's alternate standards and assessments for students with significant disabilities, including: career and life role assessment systems, extended reading, extended writing, and extended mathematics assessments. Currently, she consults with several state departments of education regarding the inclusion of students with disabilities in large-scale assessment programs including Texas, Alaska, and West Virginia. She is a member of the Washington (state) National Technical Advisory Committee (NTAC) and advises the Office of the Superintendent of Public Instruction in the area of including students with disabilities in statewide assessment systems for accountability. Dr. Almond routinely presents at national conferences regarding the inclusion of students with disabilities in large-scale assessment, including CCSSO Large Scale Assessment Conference, the annual meetings of the American Education Research Association and the National Council on Measurement in Education, and the National Meeting of the Council for Exceptional Children. She has a new chapter in *Large-Scale Assessment and Accommodations: What Works?* Edited by Cara Cahalan Laitusis and Linda L. Cook (2007) entitled "Accommodations for K-12 Standardized Assessment: Practical Implications for Policy." She coauthored a white paper entitled commissioned by the US OSEP with the Technical Work Group that is posted on the US DOE website at <http://www.osepideastthatwork.org/toolkit/index.asp>. Dr. Almond was awarded a life-time achievement award from the ARC of Oregon for service to children and adolescents with mental retardation.